SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	23 March 1998	Final REport

4. TITLE AND SUBTITLE
Model-Based 3-D Object Identification

5. FUNDING NUMBERS
DAAH04-93-G-0237

6. AUTHOR(S)
David Cyganski, R. F. Vaz, J.A. Orr

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
ECE Department
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609-7080

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 32187.4-EL

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The ATR technique developed in this project is based on a new non-linear pose estimator rather than on search mechanisms. Low false alarm rate performance is obtained by not forming a pose invariant detector but instead by incorporating pose dependent object information within the recognition process. The ATR is factored into a computationally intensive preparation process and a fast on-line target identification process. The approach is model-based and free of assumptions about the imaging process and object characteristics, and, can be applied to ATR and the estimation of pose parameters for articulated or multi-configuration targets from image and non-image sensor data.

In this work, the initial concept of the pose estimator for 1 DOF (degree-of-freedom) problems was developed into a system for N DOF whole and partially obscured target pose indexing and recognition. Performance was demonstrated at the level of filter bank implementations for 1 DOF problems at 1/17 the computational cost for unobscured targets and false alarm rates orders of magnitude better than that of the filter bank approach for obscured targets. The computational savings further increase with N for N DOF problems. The report contains ROC curves obtained from tests using the public MSTAR data set.

14. SUBJECT TERMS	15. NUMBER IF PAGES
Automatic Target Recognition Model Based ATR MSTAR INDEXING	30
	16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Enclosure 1

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

Model-based 3-D Object Identification and Pose Estimation from Linear Signal Decomposition and Direction of Arrival Analysis

David Cyganski, Richard F. Vaz, John A. Orr

March 23, 1998

U.S. Army Research Office
Grant No. DAAH 04-93-G-0237

Worcester Polytechnic Institute

Contents

List of Figures

1 Problem Statement

Model-based automatic target detection and recognition (ATD&R) systems seek to establish the presence and identity of targets from sensed data of a scene. The goal is to be able to perform this recognition task despite unknown target and/or platform position and pose; this is accomplished by the use of a target model constructed from *a priori* information about the target. The ATD&R system in some sense compares potential targets to the models of all targets of interest and determines the most likely target identity. Frequently, it is also of interest to ascertain the position and pose in 3-space of the target.

Typically, this task has been accomplished by the use of models comprised of large sets of target views, in order to allow detection of targets despite unknown position and pose. This approach, however, then requires a matching of potential targets with every image in every model, a computationally exhaustive task. The burdensome computation and enormous model sizes associated with this direct approach render such systems impractical for easily deployed, real-time ATD&R applications. Simpler "pose-invariant" models, developed by averaging over object poses, suffer from performance degradation due to the lack of target-specific information; in essence, those features which change as a target's pose is varied contain vital information about the nature of the target.

This work seeks to overcome these difficulties through means of using compact target models which, although greatly reduced in size, retain most of the essential target identity and pose information. Furthermore, these models are developed in such a way that the recognition matching procedure is direct and non-exhaustive, so that both the storage and computational requirements of the technique are greatly reduced. These dramatic benefits are achieved by a partitioning of the problem into two components: a computationally intensive, off-line model construction process, and a fast, direct on-line component which provides target identity and pose information simultaneously.

How these benefits can be gained from this partitioning can be seen from an analogy drawn to Public Key Encryption (PKE). In PKE, messages can be effortlessly decoded once the key to the code is found; this ease is due to the partitioning of the decryption into the enormously time-consuming task of finding the key and the relatively trivial decoding of the message using this key. The human cognitive system apparently knows the "key" which allows rapid object detection and pose identification from visual scenes of familiar objects.

The ATD&R system being developed under this award, based on Linear Signal Decomposition and Direction of Arrival (LSD/DOA) processing operates in the same manner: each target model is not a literal representation of the target, but rather a decryption key for images of the target. With this key, decoding of the message, *i.e.,* recognition of the target, is quickly accomplished. The burdensome task is to find the key.

The basic concept for the new technique under investigation was outlined in [1], in which recognition and pose estimation of an object were demonstrated despite a non-trivial unknown pose parameter. The distinguishing characteristics of this new approach are:

- The recognition and pose estimation procedure is fundamentally a non-linear estimator which achieves a high degree of noise immunity by not forming a pose invariant detector but instead by incorporating the object pose distribution as an integral part of the detection process.

- The recognition procedure is factored into a computationally burdensome off-line model construction process and a fast on-line image processing process.
- The usually intractable global optimization step associated with non-linear estimators is effectively and quickly solved via direction of arrival analysis, a much studied and well understood operation used in classic sonar processing.
- The construction of the object model and detection process is free of assumptions about the imaging process and object characteristics. Hence the technique can be applied to the recognition and the estimation of internal (object) and external (sensor geometry) parameters of articulated and non-physical objects.

In the original work, [1], we demonstrated that a compact object model constructed from a sufficient set of views of the object permits rapid identification of the orientation of the object through the use of Direction of Arrival (DOA) processing of the projection of new object images on the model function. The first system demonstrated a 1 degree of freedom (DOF) solution of this problem ([1]) for which the DOF is a non-trivial (not an image rotation or shift) transformation of a 3-D object.

In brief, the originally proposed effort involved: solution of the non-linear optimization problems required to demonstrate multi-DOF solutions and testing of the system with non-simulated, truthed, data so as to produce evaluation of performance in field conditions. The proposal included extending this technique to the 6 DOF case and to explore the computational techniques that will make the burdensome off-line component of the processing practical for problems with large numbers of degrees of freedom.

The initial work was conducted with FLIR imagery but during the first year we were directed by DARPA to concentrate efforts on wide area search applications with SAR imagery.

In September of 1994 DARPA introduced the additional challenge of demonstrating the improvement that LSD/DOA achieves by a direct comparison with Composite Matched Filter Banks. We were also charged with finding means to determine target pose and identity when the target is partially obscured.

In September 1995, at the ARPA ATR URI PI conference, we were informed that we would be integrated into the MSTAR effort. In particular we were asked to modify the research focus so as to produce an example of an MSTAR "Indexing" module. Hence, we modified the research focus and scheduling of our effort towards solving the problems associated with the proposed MSTAR module requirements.

As will be seen in the sections below, considerable progress has been made in all of the above areas.

2 Summary of Results

In the course of this project the investigators have made significant progress with respect to several directions concomitant to the original and revised goals and of the project. Below is a brief description of the LSD/DOA pose estimation algorithm for the purpose of introducing certain notions and terminology that will be used to summarize our results. This is followed by a list of the highlights of these research results with references to the papers and theses in which the complete derivations may be found.

The LSD/DOA technique provides a means for model-based ATR of targets which may be viewed from unknown perspectives. This technique is unique in that the models used encode and exploit the relationship between target pose and signature so that the detection process simultaneously provides both pose estimates and target identity information. These compact models incorporate the variation in target signature as a function of target pose, thus exploiting the information which is *variant* under changes in target orientation and position. Also, the target recognition process in this method is a direct computation rather than the search-based process required by many model-based ATR systems.

The LSD/DOA algorithm[1] effects a partitioning of the ATR problem into two stages: model construction and pose estimation/recognition. The model construction process involves the solution of a large, usually over determined, set of equations to determine the elements of a particular basis for the image suite. This Reciprocal Basis Set (RBS) is developed such that the pose estimation/recognition stage can be performed directly and efficiently, without searching or iteration. That is, the computational burden associated with the ATR problem is largely shifted to the model-building process in this algorithm.

Generation of the RBS target models involves a great reduction in data, as a complete suite of object views is reduced to a small set of RBS elements. The number of basis elements used, and hence the size of the target model, can be chosen according to cost/performance considerations, but is in any event very modest compared to the the data from which the model is derived.

The basis elements are generated such that linear projection of target images onto the basis elements will result in a set of inner product measures which simultaneously provide a sufficient statistic for target matching and represent the data from which target pose parameter estimates can be determined. This is due to the fact that the RBS elements are chosen to encode the target pose into these inner product measures, which are called Synthetic Wavefront Samples (SWS). These are so named because, for a given target image the RBS functions have been determined such that the SWS comprise samples of a multidimensional complex exponential wave, the direction cosines of which reveal the pose parameters of the imaged target.

A Direction of Arrival (DOA) algorithm then uses the SWS to solve for the target pose parameter estimates. If more RBS functions are used, then this larger target model allows generation of more SWS, which in turn can provide better pose estimates and more reliable target detection. The reader is referred to the original paper[1] for mathematical and implementation details of the LSD/DOA algorithm; a block diagram depicting the algorithm is given in Figure 1.

A brief list of major accomplishments produced in this project includes:

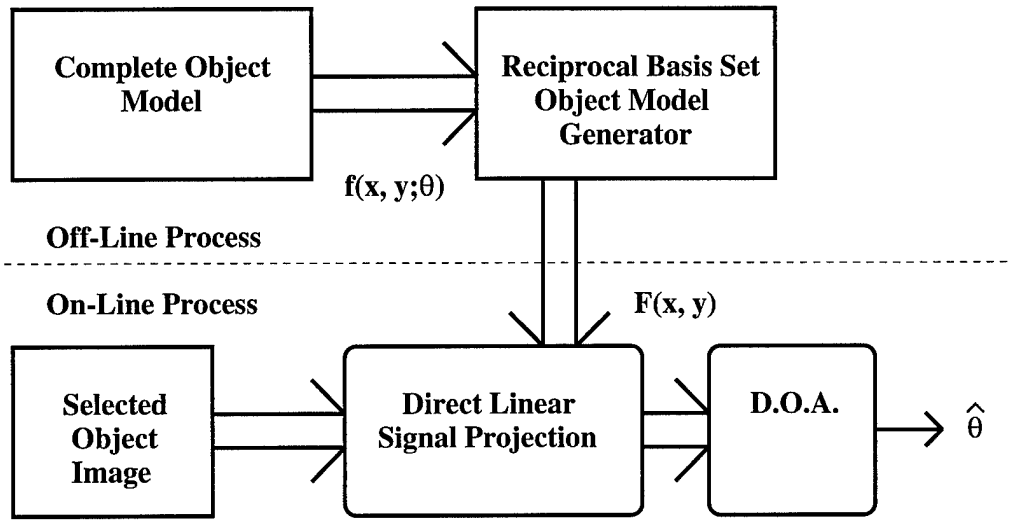

Figure 1: LSD/DOA Block Diagram

- Developed several generalizations of the non-linear signal estimation process known as Direction of Arrival (DOA) analysis. This included derivation and implementation of N-dimensional DOA systems for known non-stationary noise statistics, and, generalization of the Cramer-Rao performance bound. These generalizations were essential to the implementation of multi-DOF ATR systems [8, 10, 11, 13, 15].
- Exploration and implementation of means to parallelize the computation of large Singular Value Decompositions as arise in the off-line component of the LSD/DOA ATR [12].
- Construction of multi-DOF ATRs based upon LSD/DOA and testing with 1, 2 and 3 DOF data sets [6, 7].
- Construction of multiple target identification based on the LSD/DOA ATR [5].
- Derivation and implementation of means to construct the optimum non-stationary DOA estimator for a given target model [13].
- Implementation of a two stage ATR uses the LSD/DOA as an indexing module which yields a target pose estimate that drives a template matching system (reducing the cases that need to be tested) [14].
- Development of a generalized likelihood ratio test that permits target hypothesis testing without reference to the original target image set. This permits a very low computation and storage space implementation at a moderate cost in performance [13].
- Implementation of several testing frameworks and test data sets that permit computation of ATR ROC characteristics for large scale tests. A parallel implementation for match filter based systems permits the direct comparison of LSD/DOA and multiple matched filter based ATRs. A large 2-DOF data set was generated using the XPATCH system for testing of our first multi-DOF ATR implementations. (This test set has been replaced by the MSTAR data set and others obtained from the industry based upon field data) [7, 2].

- Theoretical development and implementation of data fusion based system for partial object recognition. The pose estimation that naturally results from the LSD/DOA process lends itself to a very computationally attractive data fusion process in which the results of several LSD/DOAs sensitive to segments of the complete target are fused into a single high performance metric [4, 3, 14, 16].

- Evaluation of the LSD/DOA performance with respect to the public MSTAR data set. (In the course of this evaluation several new methods were developed for tuning LSD/DOA to the characteristics of target sets and clutter environments) [2].

The complete text of all references connected with the above results are all available on the World-Wide Web at:
`http://xfactor.wpi.edu/RecentWork.html`

The following subsections summarize the overall systems that resulted from the above research in the context of the most recently obtained performance results. The first summarizes the approach that was used and results obtained for the case of whole target recognition from SAR images in the MSTAR data set. The second summarizes the means and results of our investigation of partial target ATR based on the information fusion made possible by the pose information derived from the LSD/DOA process.

3 Initial MSTAR Performance Results

One of the objectives of the final stages of this work is to establish the performance of the LSD/DOA ATR algorithm. The testbed against which performance is to be ascertained was specific by DARPA as being derived from the high resolution Synthetic Aperture Radar (SAR) data collected as part of the DARPA/WL Moving and Stationary Target Acquisition and Recognition (MSTAR) program.

3.1 The System Under Test

The ATR under test is an indexed template matching variation of the LSD/DOA ATR algorithm developed in the course of this project in response to DARPA's request for a demonstration of indexing accomplished by application of the LSD/DOA pose estimation capability.

Schematically, the algorithm may be decomposed into three stages as shown in Figure 2. This algorithm has been applied to detection and classification problems utilizing sensor modalities other than SAR, and has been adapted to accommodate partially occluded targets [4, 3]. The Linear Signal Decomposition/Direction-of-Arrival (LSD/DOA) pose estimator is the signature stage of this ATR, and serves as an indexing module into the template matcher. Although the LSD/DOA pose estimation stage has been coupled successfully with other non-template matching target discrimination stages [1, 9], it is the template matching configuration shown in Figure 2 that is the concern of this performance analysis. A brief description of each processing stage is presented below.

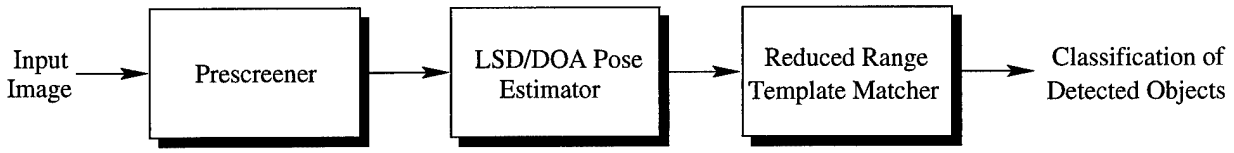

Figure 2: The ATR comprises three-stages as shown.

3.2 The Prescreener

The prescreener is adapted from a description of the prescreening stage of a three-stage ATR system implemented by Lincoln Labs [18]. It is a two parameter CFAR detector, which locates candidate regions in an image that may contain targets by searching for pixels in a SAR image that lie on the upper tail of the clutter distribution such that

$$\frac{x_t - \mu_c}{\sigma_c} \underset{H_1}{\overset{H_0}{<>}} \tau_{CFAR}. \tag{1}$$

where μ_c and σ_c are the mean and standard deviation estimations of the clutter. For the tests presented in this paper, the CFAR threshold, τ_{CFAR}, is set such that every target instance in clutter from both the test and training data sets for the ATR satisfies the lower inequality of the expression above. Thus the prescreener can never adversely affect the probability of detection for the ATR.

3.3 Brief Review of the LSD/DOA Pose Estimator

We recall that any vector has a unique expansion relative to some basis, and that there exists a dual (reciprocal) basis which can facilitate recovery of the expansion coefficients of that vector. Consider a vector A in terms of a basis $B_1 \ldots B_n$ of the form

$$A = \sum_n B_n e^{ik_n\xi} \tag{2}$$

where the complex exponentials are the coefficients of the vector A. The dual basis, $B'_1 \ldots B'_n$, of $B_1 \ldots B_n$ is such that

$$< B'_n, A >= e^{ik_n\xi} \tag{3}$$

The form for A is intentionally suggestive. If an image point is Fourier expanded in terms of a pose parameter, ξ, then a dual basis obtained from the Fourier coefficients (corresponding to the basis B) could be used to recover the complex exponentials corresponding in n-dimensions to plane waves with direction cosines ξ and sampled on a lattice defined by the k_n. The recovery of direction cosines from wavefront samples in noise corresponds to the nonlinear direction-of-arrival estimation problem. The current implementation of the LSD/DOA module uses a direction-of- arrival estimator based on components originally derived by Kay [19] and Lank *et al.*, [20] and subsequently improved by Cyganski and Fraser [21].

The generation of a reciprocal basis set (RBS) for each target is part of the training process for the ATR. RBS generation is computationally intensive, but is conducted off-line. Projection of a test image onto an RBS is substantially more computationally efficient. The ultimate result of this processing stage is an index into the template library with the potential to substantially reduce the volume of pose space that must be searched by the template matcher.

3.4 The Reduced Range Template Matcher

The template matcher, from Figure 2, produces a sum-of-squared-differences match metric between template image and the current test image window. A set of representations of a target sampled over some parameter vector, such as pose, comprises the template library for the target. The template library can be considered as a set of images sampled over pose space. The LSD/DOA pose estimation stage, discussed above, provides an index, k, (position vector) into the pose space. Viable candidate templates are those within a specified distance, ρ_k, from the position, k, estimated by the LSD/DOA processing stage. Thus, there is a subset B_k with elements that satisfy the distance criterion given above, which becomes the searched part of the template library. The matching process between a test image, r, and templates t^i, might be written

$$\min_i \sum_{j \in M_i} (r_j - t_j^i)^2 \text{where, } t^i \in B_k, \tag{4}$$

where the summation is over the pixel locations contained within a pose dependent masked region, M_i, corresponding to the i^{th} template image. The minimization is over only that portion of the template library $t^i \in B_k$.

The combination of the LSD/DOA pose estimation stage with the template matcher is illustrated in Figure 3. The distance ρ_k is a design parameter obtained during the training phase of the ATR.

3.5 Test Data Description

The MSTAR program is directed toward the development of the next generation model-based ATR. In support of the program, a substantial data collection effort was undertaken with a state-of-the-art SAR sensor and consists of X-band SAR imagery with 1 foot by 1 foot resolution. The data sets contain a large number of target types, some in multiple versions and/or configurations, well-sampled over azimuth and depression angle. Clutter scenes also vary widely from woodlands to tilled fields to urban scenes. The target and clutter data used in this study are taken from the MSTAR Public Release Data set.

MSTAR target chips are 128 by 128 pixels in size. For our purposes, the template library was formed by cutting these down to 48 by 48 pixel target chips and transforming to a dB scale. A standard median filter and a variant which excludes all pixels brighter than some threshold were both applied to dB-scaled imagery. Figure 4 illustrates the preprocessing steps applied to the MSTAR imagery before submitting it to the ATR.

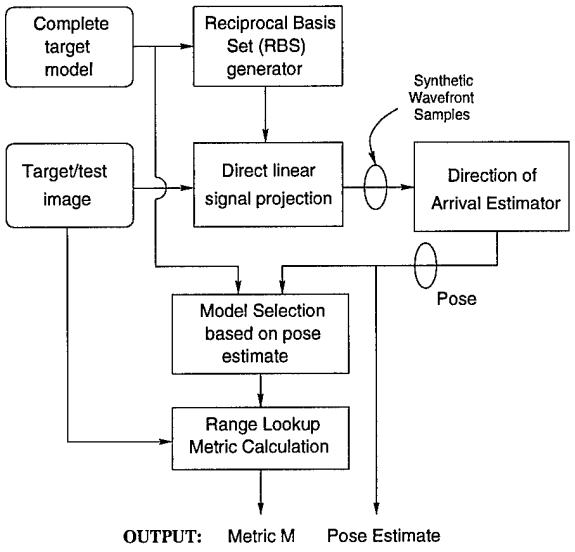

Figure 3: A more detailed representation of the pose estimation and template matching modules.

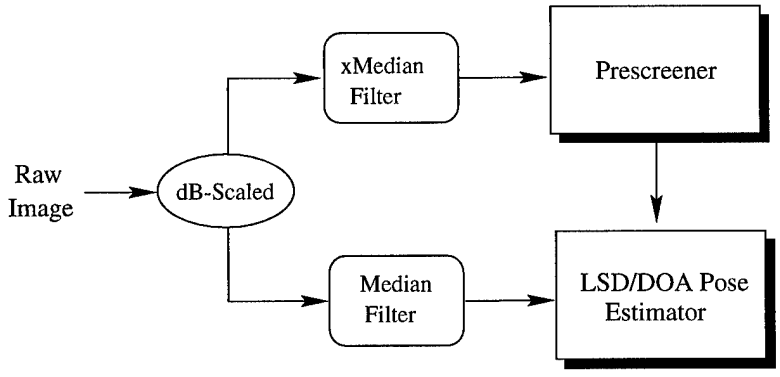

Figure 4: Input configuration illustrating the operations, including two types of median filters, used to preprocess the MSTAR imagery.

3.6 Subsystem Performance

Much of the performance discussion presented here is in the context of binary hypotheses: "No target is present" (H0) and "A target is present" (H1). The test item for all of the subsystem tests of the ATR is a BMP-2, sn6563, from the 17 degree depression angle imagery. Pose variation is restricted to one degree-of-freedom, azimuth. A single clutter image of a rural setting with trees (hb06158) and acquired at a depression angle of 15 degrees is used throughout this section on subsystem performance, except where noted otherwise.

3.6.1 The Prescreener

The success of the prescreening process is at one level measured by the percentage of the clutter image that is screened out. Using both the training and test target suites to set the CFAR threshold so no H1 cases are screened out, it was observed on a variety of the clutter images that the prescreener processing stage removed 99% (hb06160, urban scene) to 99.99% (hb06172, tilled fields-no trees) of the background. Obviously, the computational load on subsequent system calculations for H0 cases is markedly reduced by the prescreener.

3.6.2 LSD/DOA Pose Estimation

Pose estimation error can be simply defined as the difference between actual and estimated pose values observed for a given target. For the BMP-2 against rural clutter, the pose estimation error is shown in Figure 5. The error distribution in Figure 5 can be used to select a value for the distance ρ, introduced in the discussion of the template matcher above. For example, a value of $\rho = 5$ corresponds to a search range through the template library of ± 5 indices around the index, k, provided by the LSD/DOA module. If there are 44 images in the template library, then the searchable portion of the library consists of 11 images, which is a 75% reduction in the volume of the template library which must be searched to find the best match.

3.6.3 Template Matching

The sources of discrepancy that degrade template matching performance include registration error between test target and template, speckle and noise processes, discrepancy in target configuration between test target and template, discrepancy in versions between test target and template, template density in the template library, and last but not least, a range ρ around a pose index from the LSD/DOA module that does not include the actual solution (an out-of-bounds error). To remove configuration and version discrepancies, we tested on target images taken from the same set as, but disjoint from, those used to produce the template library. Thus registration and speckle remain as the two primary culprits affecting performance.

As noted above, the LSD/DOA processing stage worked equally well with linearly scaled and dB scaled imagery. However, the template matcher is very sensitive to the scaling used and performed dramatically better with dB scaled imagery. This situation is not surprising given that logarithmic scaling turns multiplicative speckle into an additive process reasonably approximated as a Gaussian process [18].

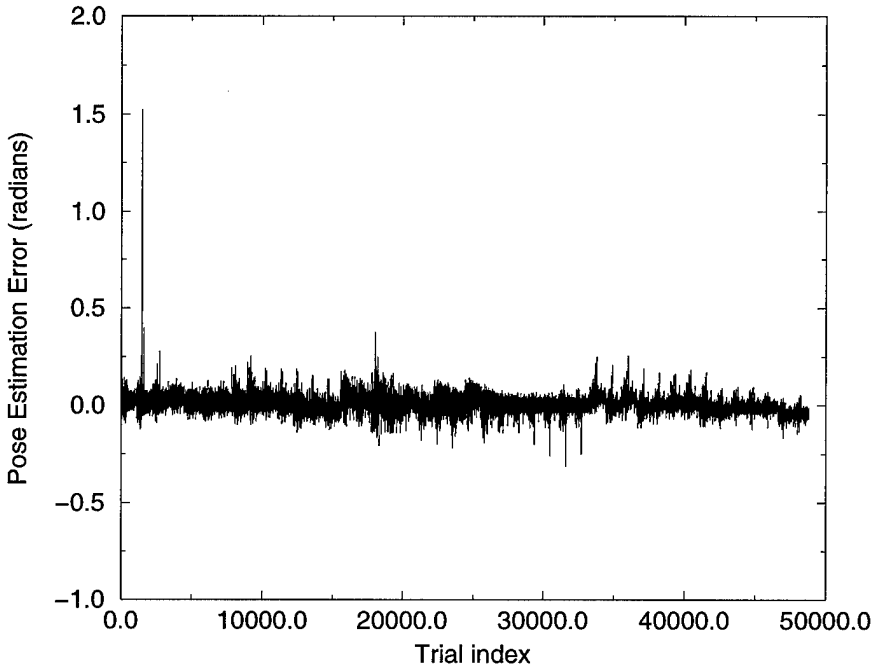

Figure 5: The discrepancy between actual and estimated azimuth angles as target pose and clutter image are varied.

3.7 Performance Evaluation Procedure

Perhaps the most common vehicle for expressing ATR performance, the receiver operating characteristic (ROC) diagram, plots the probability of detection against the probability of false alarm, or equivalently, against the false alarm rate. The ROC curve is typically generated by parametrically varying the boundary between the acceptance and rejection regions of the observation space and computing the pair of probabilities at each boundary value.

The basic procedure employed in this evaluation to generate H1 cases is to overlay target pixels from a target chip onto a 48×48 pixel window over a clutter image. A 48×48 window is obtained from every possible position in the clutter image, producing both H0 and H1 cases for test at every scanable pixel in the clutter image. The advantage of this approach is that it makes maximum use of the clutter data available. Implicit in this approach is the assumption that each template target and each test target overlay are both centered in the window for each H1 trial.

The training image set, used to construct the RBS for the LSD/DOA module and the template library, contains image chips of the BMP-2, sn6563, sampled at azimuth angles

$$\begin{array}{rrrrrrr} 0.49 & 3.49 & 7.49 & 13.49 & 15.49 & 18.49 & 20.49 \\ 24.49 & 26.49 & 28.49 & 30.49 & 33.49 & 35.49 & 37.49 \\ 39.49 & 43.49 & 45.49 & 47.49 & 49.49 & 51.49 & 55.49 \\ 59.49 & 62.49 & 64.49 & 67.49 & 69.49 & 73.49 & 76.49 \\ 79.49 & 82.49 & 84.49 & 88.49 & 90.49 & 92.49 & 94.49 \\ 96.49 & 98.49 & 101.49 & 105.49 & 107.49 & 110.49 & 112.49 \\ 115.49 & 118.49 & & & & & \end{array} \tag{5}$$

in degrees. The set of test cases was sampled for the same target at azimuth angles

$$\begin{array}{rrrrrrr} 2.49 & 4.49 & 12.49 & 14.49 & 17.49 & 19.49 & 23.49 \\ 25.49 & 27.49 & 29.49 & 32.49 & 34.49 & 36.49 & 38.49 \\ 42.49 & 44.49 & 46.49 & 48.49 & 50.49 & 54.49 & 58.49 \\ 60.49 & 63.49 & 65.49 & 68.49 & 70.49 & 74.49 & 78.49 \\ 81.49 & 83.49 & 87.49 & 89.49 & 91.49 & 93.49 & 95.49 \\ 97.49 & 100.49 & 102.49 & 106.49 & 108.49 & 111.49 & 113.49 \\ 116.49 & 119.49 & & & & & \end{array} \tag{6}$$

Note the two sets of images are disjoint. The target imagery corresponds to a 17 degree depression angle and is overlayed onto clutter imagery acquired at 15 degrees depression angle.

The objective of this set of ROC curve generation tests is to establish a performance baseline on this target against which subsequent tests may be compared: especially tests that specifically incorporate Extended Operating Conditions (EOCs) [22, 23] into their design.

3.8 Test Results

Figure 6 shows the ROC curves for the BMP-2 against a rural scene and an urban scene obtained from a single 0.1 km^2 clutter image. By exploiting all clutter imagery at 17 degrees

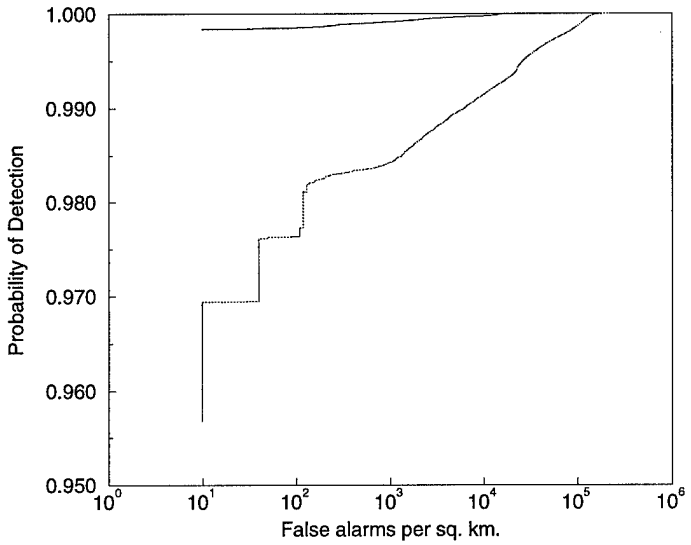

Figure 6: ROC diagram for detection of a BMP-2 in patches of rural (solid-black) and urban (dotted-red) clutter.

depression angle available in the public MSTAR data set (a total of around 10 km^2), we obtain an ROC describing detection performance at false alarm rates as low as 0.1 false alarms per km^2, which is shown in Figure 7.

4 Partially Obscured Target Recognition

The fundamental problem in model-based ATR systems is that the targets to be detected have an appearance which is a function of the location and orientation of the target with respect to the sensing device(s). This problem is further complicated by the addition of partial target occlusion.

The simplest solution to partial obscuration with a model-based approach is to train the obscuration into the image model. This simple method, however, is usually unusable due to the large complexity associated with the process of searching or training over the additional degrees of freedom introduced by various degrees and modes of obscuration. Some methods that have been implemented apply accelerated global search techniques in an attempt to overcome this growth in complexity. Other systems allow partial feature set matches found by use of rule based processing.

Another common approach for dealing with the detection of partially occluded targets is the use of a match metric with appropriate behavior with respect to occlusion. These metrics allow, and in some cases quantify, deviations of portions of the input image from the target model due to partial obscuration.

The technique explored in this part of our work, which we have named Partial Evidence

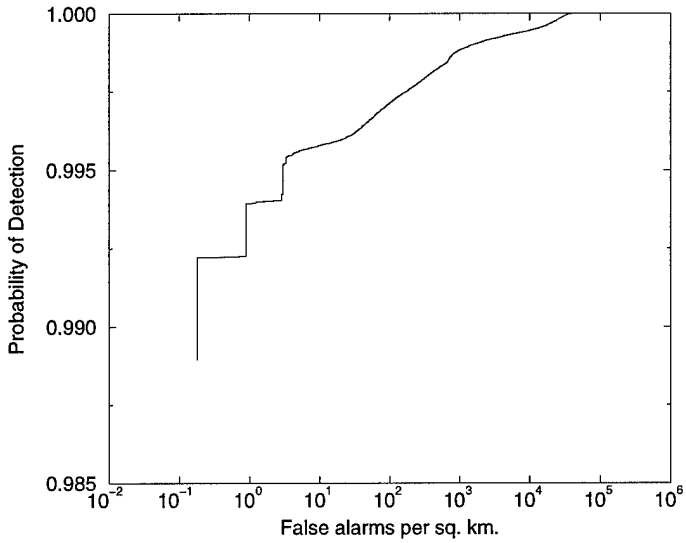

Figure 7: ROC diagram detection of a BMP-2 over extended area (10 km^2) containing both rural and urban scenes.

Reconstruction From Object Restricted Measures (PERFORM), obtains several match and pose values from subsystems that seek to identify portions of the object. The LSD/DOA based subsystems are themselves capable of finding the identity and pose of the components of the object. The object pose information is instrumental in fusing the several metrics into a single, occlusion robust, whole target metric.

4.1 Development of PERFORM

During our investigations into the behavior of the single RBS based LSD/DOA ATR system we determined that a root cause for performance loss in systems with increasingly "detailed" object models (contrary to the expected trend) is a concentration of ATR system reliance on pixels in the outer most region of target support over all poses. That is, the ATR algorithm attempts to glean its largest share of pose related information from the very pixels that are most likely to be corrupted by clutter.

The solution that we originally applied involved use of a weighting of the pixels in the region of interest (ROI) in accordance with the probability that the pixel contains target information and not clutter. We found that a graded weighting yields far better results than both no weighting or the other extreme of complete deletion of these pixels (since they do contain valuable target pose information at times.)

Ultimately, the best performance for LSD/DOA would be obtained for the case wherein the target always occupied all pixels within the convex hull of the target support regions over all poses. If instead, we restrict the filter support region to occupy only the area inside the intersection of all target support regions over all target poses, then the loss of target

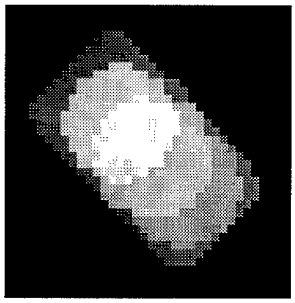

Figure 8: Three embedded cover filter support regions shown superimposed on one pose of a T72 tank target.

information causes a loss of pose estimation accuracy that offsets the benefit.

The investigation resulted in the implementation of an approach that eliminates the effects of surrounding clutter while utilizing most of the pose information throughout the target support region. This method employs several LSD/DOA recognition systems which focus on well chosen subregions of the target support area. The resulting partial support evidence is accrued as a set of functions, that resemble conditional probability functions, to form a single pose estimate and single hypothesis metric. The advantages of this approach are summarized below:

- By confining an RBS support region to the interior of the object support intersection over the full pose range, we can obtain an ideal elimination of clutter effects with an accompanying loss of object information and hence discrimination power.
- By using several such embedded "cover" estimators we can develop a set of partial object evidence maps which can be assembled to produce a single object pose and identity hypothesis.
- One obtains greatly increased robustness to object obscuration in so much as certain cover estimators may be altogether unaffected by partial obscuration.

For the purposes of the first implementation we utilized just three embedded covers. Figure 8 depicts the three covers associated with a Soviet T72 tank target. These three covers are then used in the system described by Figure 9.

In effect, each instance of the algorithm, that is, each application of a "cover" filter followed by the LSD/DOA ATR process, results in match metric and estimated pose information for a different section of the target. Each such pair can be represented as a single complex valued metric with phase indicating pose. The set of such metrics for each position of a cover filter forms a complex valued metric function for each RBS function known as the single cover complex metric function (SCCMF) as shown in Figure 10.

The pose information in these functions is used to re-map the various match metric values to appropriate locations in a fused complex match metric information structure. The pose estimates themselves are "coherently" added so as to cause re-enforcement in the case of true targets and destructive interference otherwise. The warping function applied is based upon a vector flow field defined by the phase of the original complex metric value and a displacement value associated with that pose during the construction of the cover filters.

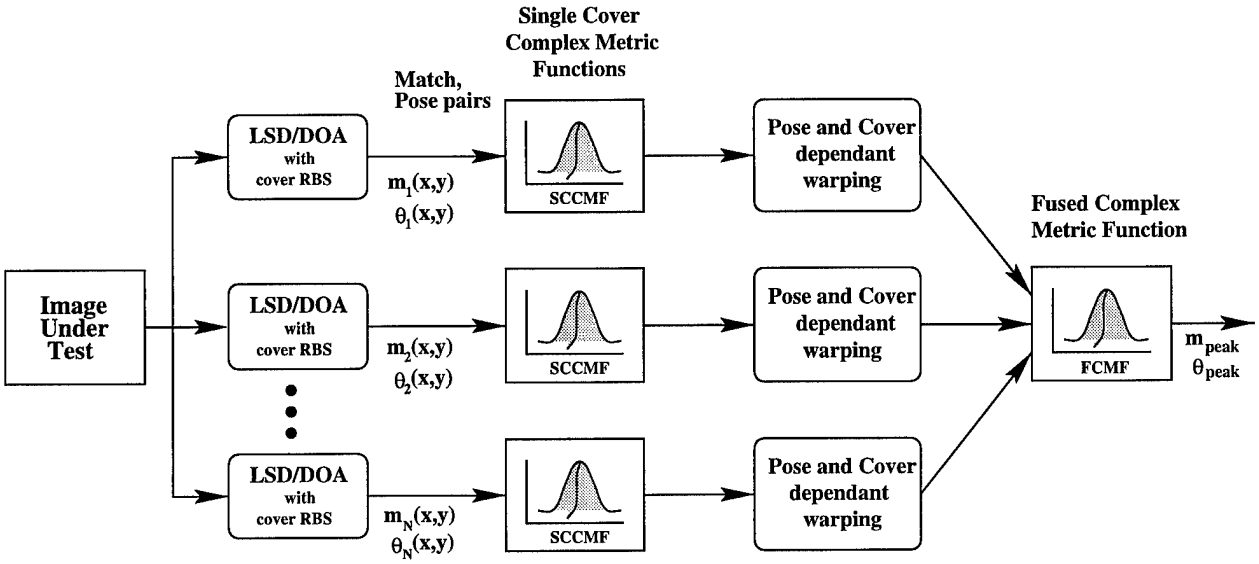

Figure 9: PERFORM system block diagram.

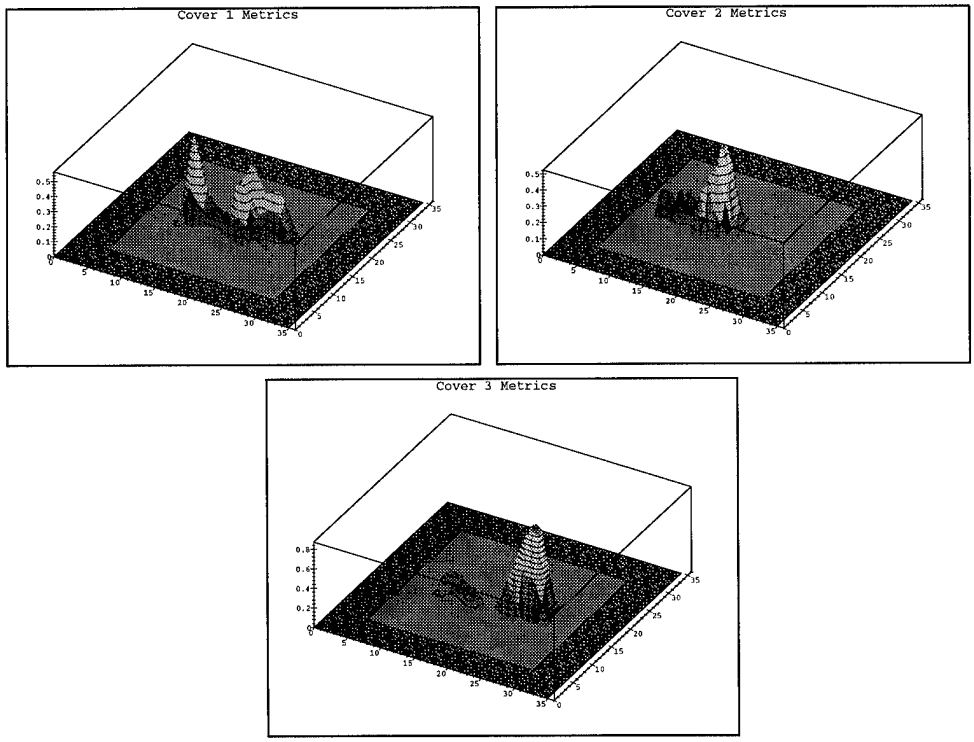

Figure 10: Magnitude of the SCCMF for each of the three covers in the example.

A merged version of the three cover responses, for which the warping operation was not applied, can be found on the top of Figure 11. As can be seen, there are three distinct peaks representing the peak response of the three covers. Also shown in Figure 11 is the final fused metric image or fused complex metric function (FCMF). The maximum response corresponding to the true center of the target is now clearly in evidence.

Since the PERFORM method uses partial evidence to recognize a target, it should be well suited for the recognition of partially obscured targets. Figure 12 depicts an artificially speckled unobscured target and a partially obscured version of that target. By looking at the merged and fused images produced by the PERFORM method for these two targets, one can see the significance of the PERFORM method.

Figure 13 shows the merged and fused metrics for the unobscured and obscured target respectively. As was the case in our other unobscured target example, the unobscured merged metric contains three distinct peaks which represent the centers of each of the three covers. Notice, however, that the obscured merged metric only contains two distinct peaks. This is due to the occlusion of most of the target associated with the third cover.

As one can see in both of the fused metric images there is a large peak associated with the target center. The only difference in the metrics is that the obscured metric is slightly smaller due to the loss of information in the obscuration. Thus, the addition of partial obscuration does not make it impossible to detect the target but simply reduces the level of the match metric.

4.2 Improved Metric Fusion Process

After obtaining initially encouraging results [6], additional research was applied towards improving the PERFORM metric fusion process rather than that of the individual LSD/DOA subsystems. This goal was been pursued through construction of a Constant False Alarm Rate (CFAR) detector based on probabilistic analysis of the output of the LSD/DOA subsystems.

As before, the front end of the PERFORM ATR consists of three LSD/DOA subsystems. An image under test is fed into each of them producing a map of metrics and corresponding poses representing responses to a given cover RBS. Subsequently, all three metric/pose maps are processed by several stages of the algorithm with the purpose of obtaining the final, fused, match measure as illustrated in the flow diagram of Figure 14.

As stated above, metrics in all of the maps are shifted by a non-linear transformation which is driven by the pose values associated with each metric in the metric/pose pair map, to a single point of reference for a given target. In the case of a T72 tank example each of the metrics produced by the front and rear covers of the tank are shifted to the apparent center of the tank according to their original locations and respective pose estimates.

Even if a combination of the metrics would merge into a single point, however, the combination will still be rejected by this implementation of the system if their original locations indicate that it would be impossible for the pair to simultaneously correspond to a pose driven transformation of the rigid geometry of a tank.

In the following stage, a set of likelihood ratio tests is performed in order to decide how many of the target covers are occupied by a corresponding part of a target in the test scene. This is not undertaken in an effort to make a final decision about the presence or absence of

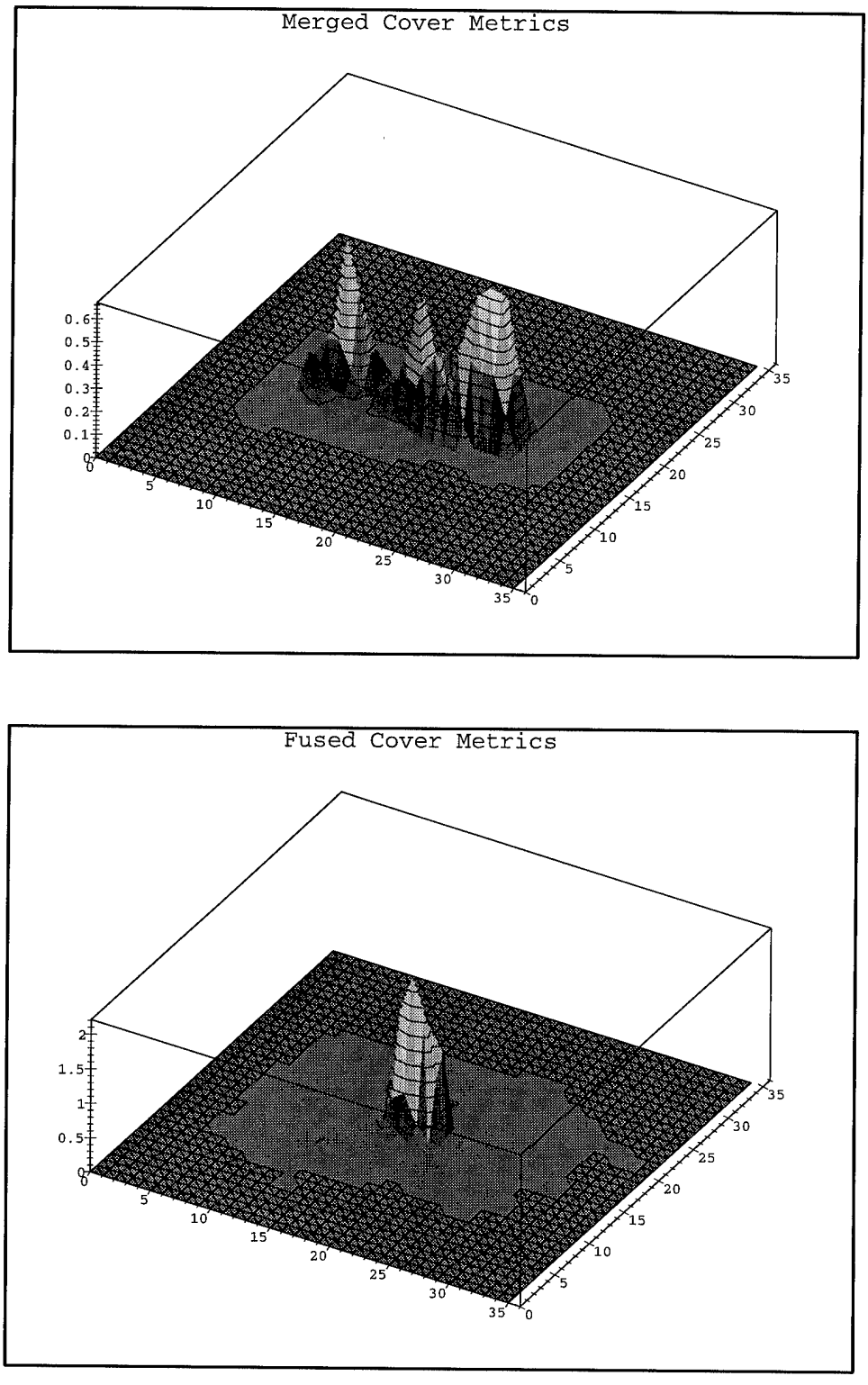

Figure 11: Merged metrics (top) and fused metrics (bottom) from 3 SCCMF in the example.

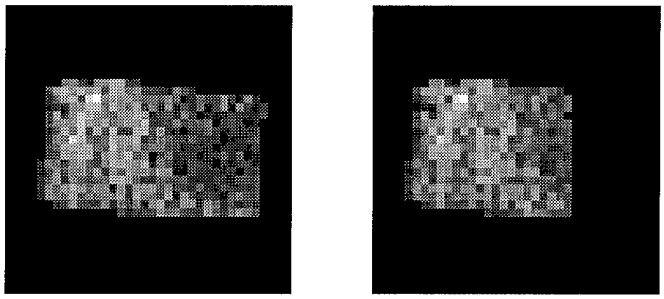

Figure 12: Example unobscured and obscured target exemplars.

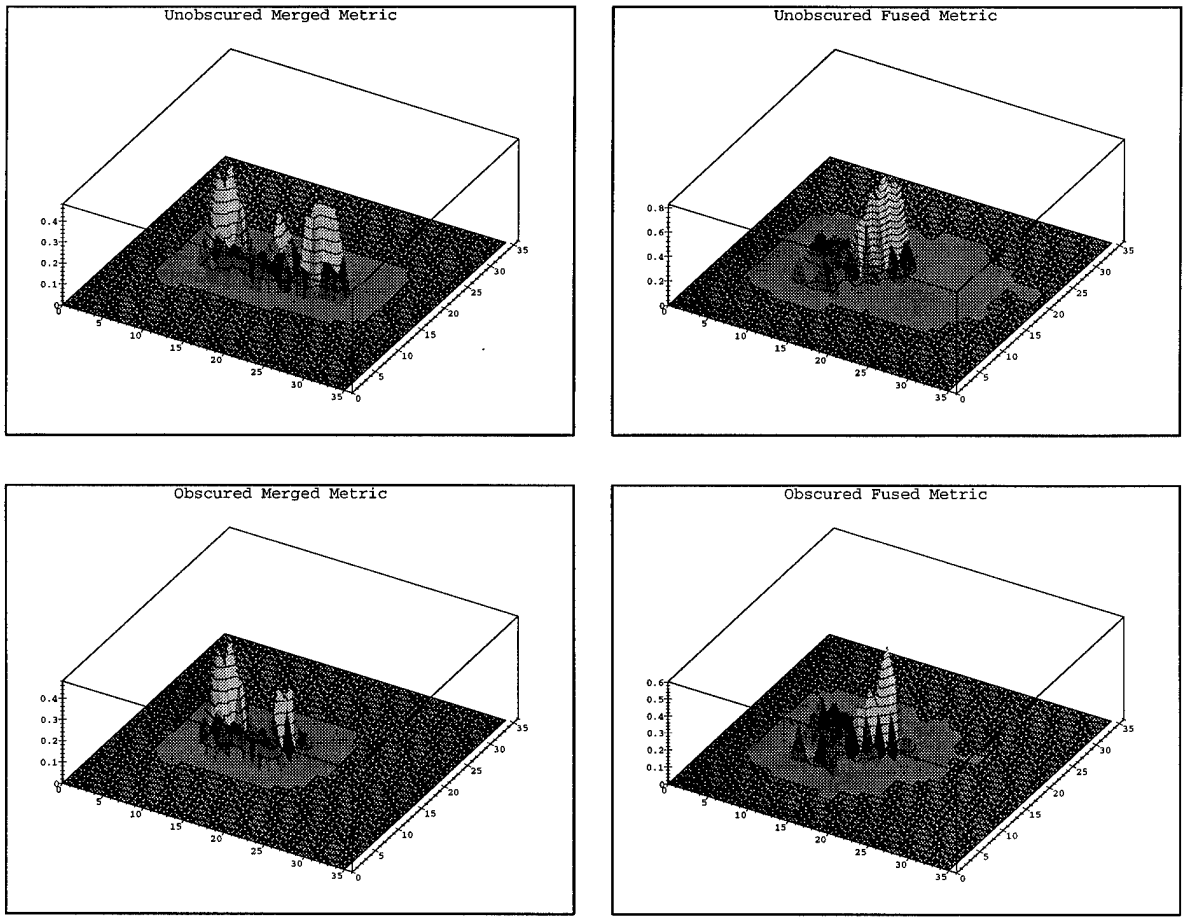

Figure 13: Merged and fused metrics for unobscured (top) and obscured (bottom) examples

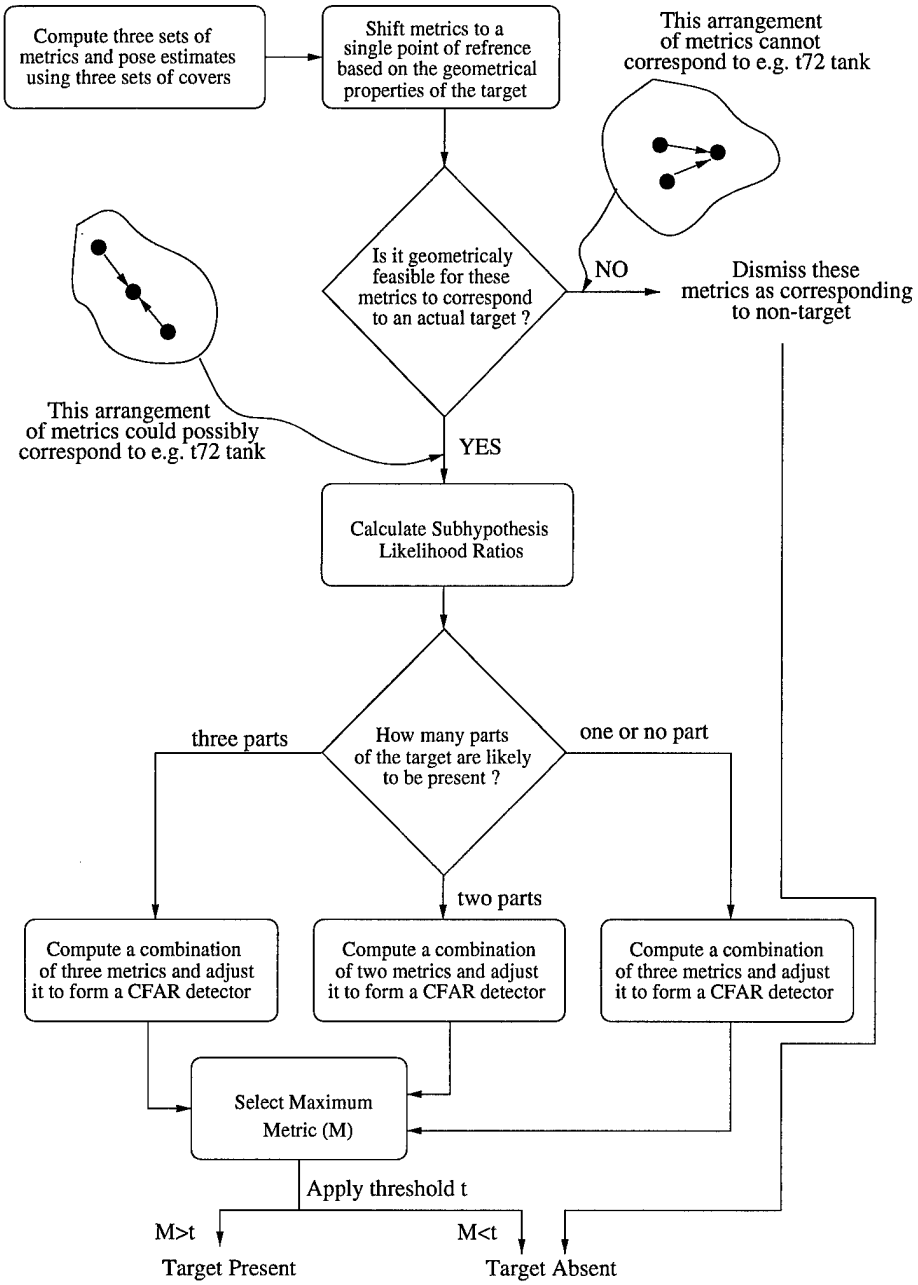

Figure 14: PERFORM metric fusion process.

the target at the current location, but rather in an attempt to remove from consideration all the metrics generated by applications of cover filters which could not possibly support the presence of an object within that cover region. The decision is based on an assessment of clutter content in the image under test by analysis of the values of the metrics obtained in relationship to statistical measures of the image in the cover surround. Only cases where at least two parts of the target are present are of concern in this particular implementation.

Through theoretical analysis of the subsystem metrics it was determined that the optimum decision metric is formed as a linear combination the subsystem metrics with appropriate weighting factors. Thus one forms the linear combination of three metrics in the case of possibly three target parts present and the linear combination of just two metrics in the case of possibly two target parts present.

As can be expected, interpreting both such metrics in the same way would be inappropriate. The final metric, which is a composition of three subsystem metrics must be judged on a different basis than the two-subsystem metric based result. Hence there is a need for a Constant False Alarm Rate (CFAR) detector, which will establish a direct relationship between decision thresholds used with two and three cover metrics.

Such a detector can be constructed analytically by computing the probabilities of false alarm rate for both cases, equating them and solving for the desired threshold. As a result two thresholds, one for the two and one for the three cover metrics, need to be calculated in order to make the final decision as to whether the given target is present or missing.

A more detailed description of this analytical approach and its results are further discussed below.

4.3 PERFORM Evaluation Procedures, Results and Conclusions

Although PERFORM can be used as a single stage ATR system, use of an additional stage such as the pre-screener can greatly reduce overall computational requirements.

In one implementation of an ATR system by Lincoln Labs, a pre-screener is used as the first stage of a three-stage SAR ATR system. It is a two parameter CFAR detector, which locates candidates for targets in the image under test by searching for high amplitude pixels in a SAR image. In order to minimize the computational requirements of this stage, reduced resolution images are processed. In their case, the pre-screener works on one meter resolution images instead of the full one foot resolution imagery. Having a pre-screener at the first stage of a PERFORM based system greatly reduces the number of images that have to enter the final, computationally more intensive, classification stage. PERFORM as augmented by the pre-screener is shown in Figure 15.

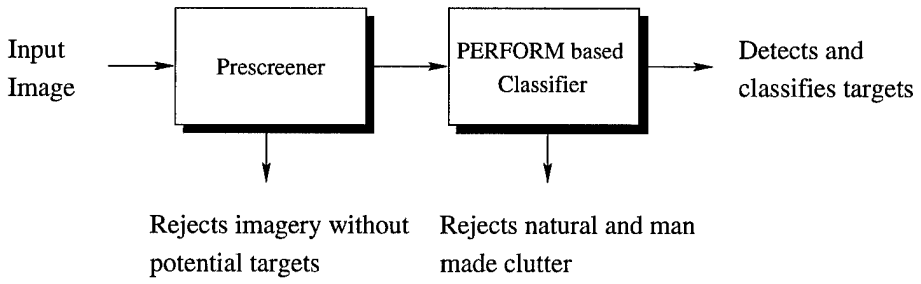

Figure 15: Complete PERFORM based ATR system.

The system as depicted in Figure 15 has been implemented and tested using SAR imagery. The target data was generated from spotlight SAR phase history files provided by

Wright Laboratories, Wright-Patterson AFB. The test results presented below are based on data of a Soviet T72 tank. The target exemplars were generated from L band data, a 10 degree elevation angle, and HH and VV polarization data which were used to form a single, polarimetrically whitened image.

The background data was obtained from Lincoln Laboratories (ADTS data set). The images used were polarimetrically whitened SAR images depicting terrain in Stockbridge NY at 1 ft. by 1 ft. resolution. The final test images were obtained by overlaying the targets onto the clutter backgrounds by masking out a region of the clutter corresponding to the convex hull of the brightest target pixels and inserting the target image into the masked area.

Typically, simulating realistic obscuration (e.g. SAR layover) is not an easy task. For these initial tests, a simplified approach was taken. Portions of the target's pixels were replaced with clutter background on one side of a line passing through the target opposite the direction of illumination. The obscured target suite thus generated contained speckled target images with obscuration ranging from 5 to 25 percent of the target energy.

Tests, using the same data, were performed for the LSD/DOA Range Lookup (see MSTAR evaluation description above), Composite Matched Filter and PERFORM ATRs. The results of these tests are presented as Receiver Operating Characteristic (ROC) curves in Figure 16 and Figure 17. Also, the ROC of the first implementation of PERFORM, which did not use the geometric constraints and the composite hypothesis based CFAR detection system, is shown on the same graph [6].

Applying the additional decision process to determining the occupancy of the PERFORM test regions, and applying the CFAR based final decision performance was improved significantly over the initial PERFORM implementation. When tested with unobscured targets, the current version of PERFORM also rivals the performance of Range Lookup LSD/DOA and CMF based ATRs, while significantly outperforming these two systems when exposed to obscured targets.

5 Conclusions

In the course of this project we have taken the initial concept of a 1 DOF (degree-of-freedom) model based pose estimator based upon the computationally fast and search-mechanism free method first identified by the PIs and developed several end-to-end ATR systems. The following list summarizes the most important outcomes of this effort, the details of which can be found in the interim reports and publications related to this work:

- Development and implementation of the non-linear estimator theory and algorithms needed to solve N-DOF ATR problems.
- Demonstration of ATR systems for whole-target on-clutter recognition that have been tested with 1, 2, and 3-DOF targets and imagery (including SAR, FLIR and Optical image sets.)
- Test results indicate that while constant-level-of-performance N-DOF matched filter bank approaches scale as K^N for some constant K, the N-DOF systems based on this

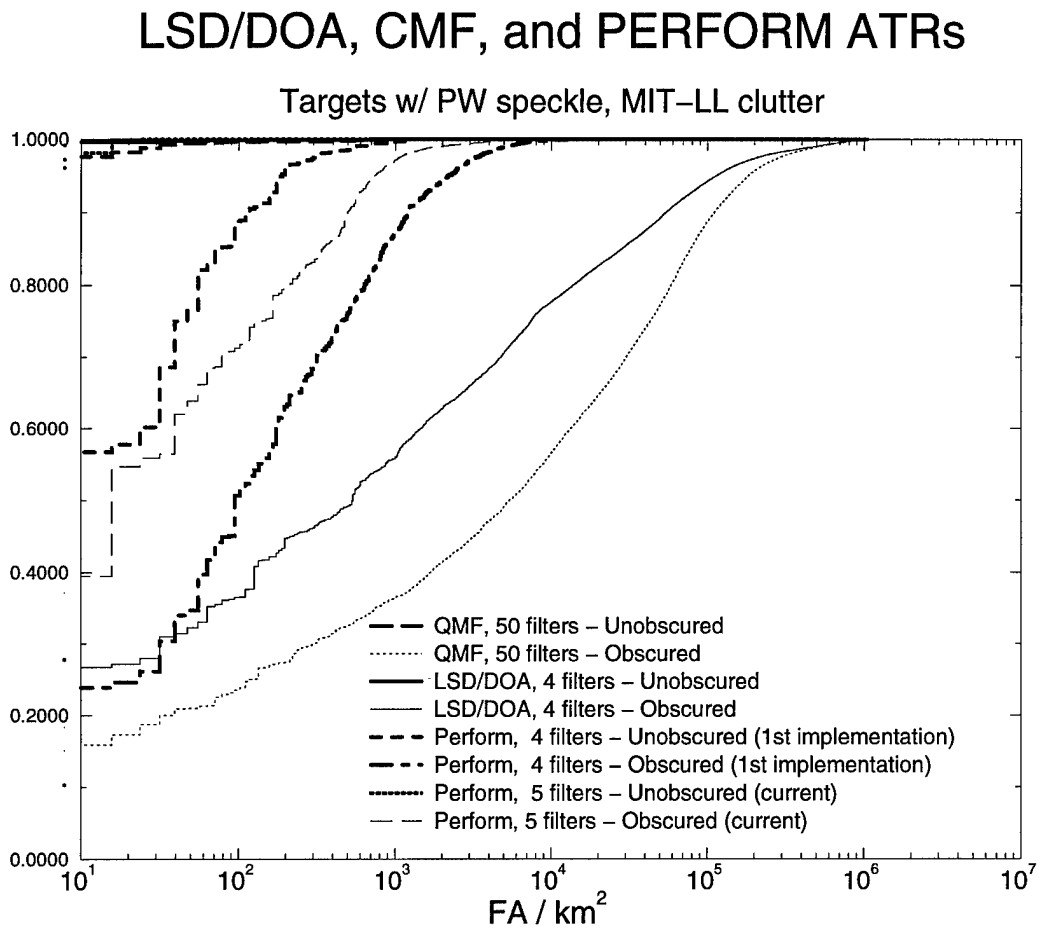

Figure 16: Comparison between CMF, Range Lookup LSD/DOA and PERFORM systems.

approach present computational complexity of the form $(K/D)^N$ where D is a target dependent factor that may be as high as 20. Thus, where computational complexity still grows exponentially with increasing degrees of freedom, the base of this exponential is much lower and the advantage (expressed as a ratio) thus also grows exponentially.

- The pose parameter information can be used as an indexing value to drive a limited-search model-matching system. This pose estimator is quite robust and has been used to demonstrate performance equivalent to that of an exhaustive search based system with $\frac{1}{10}$th or fewer search trials. This approach was used to obtain the preliminary MSTAR based performance results shown above.

- The pose estimation that is inherent to this ATR approach allows construction of a multi-look information fusion system that can be also be used for partial object recognition (as detailed above).

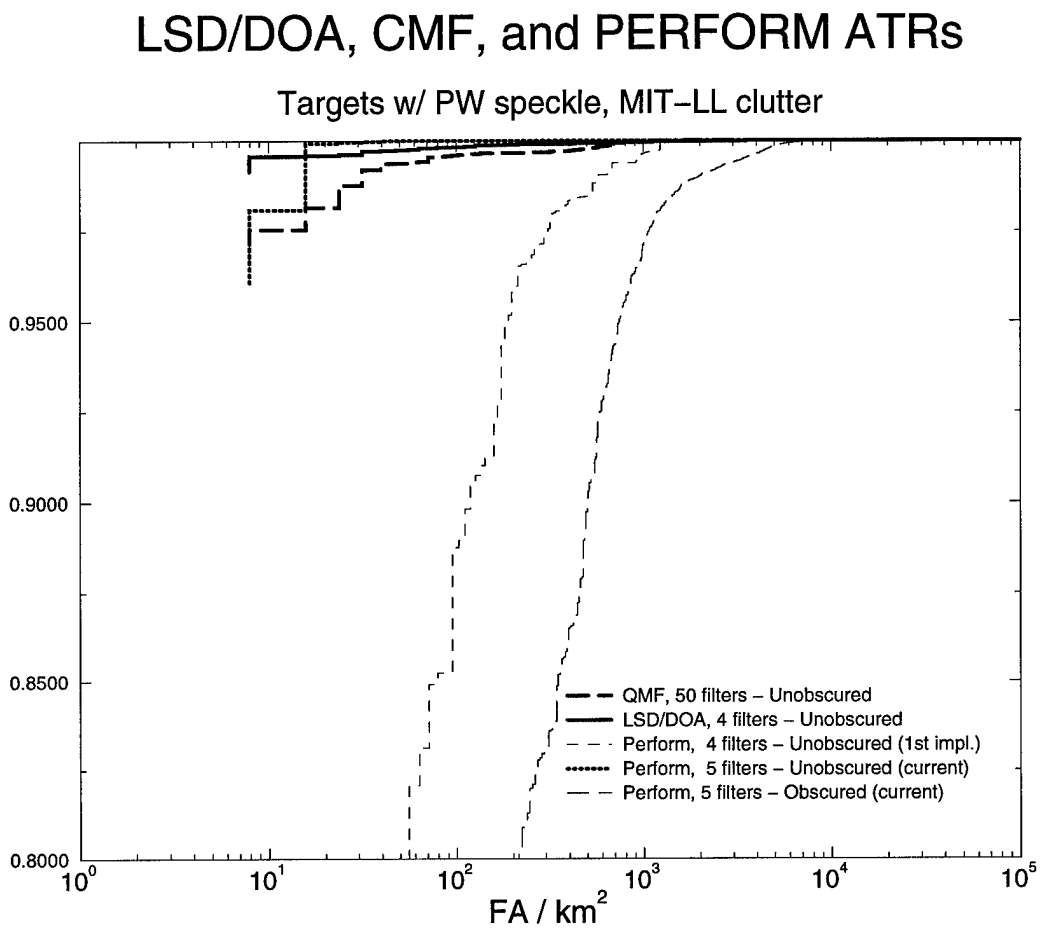

Figure 17: Expanded view of ROC performance graph in previous figure.

We have both undertaken and are currently engaged in several projects based upon further extension, testing and technology transfer of the ATR techniques that have arisen from this work.

6 List of Publications

1. David Cyganski, James Kilian, Deborah Fraser, "Baseline Performance Analysis of the LSD/DOA ATR Against MSTAR Data", SPIE AeroSense '98, Conference 3370, Algorithms for Synthetic Aperture Radar Imagery V, April 1998, Orlando, Florida.

2. Witold Jachimczyk, David Cyganski , "Enhancements of Pose-tagged Partial Evidence Fusion SAR ATR," AeroSense '97, Conference 3070, Algorithms for Synthetic Aperture Radar Imagery IV, April 1997, Orlando, Florida.

3. B.K. Hill, D. Cyganski, R.F. Vaz , "SAR ATR via pose-tagged partial evidence fusion," AeroSense '96, Conf. 2757, Algorithms for Synthetic Aperture Radar Imagery III, April 1996, Orlando, Fla.

4. Hill, B.K., Cyganski, D., R.F. Vaz, "Multi-Target Discrimination with Linear Signal Decomposition/Direction of Arrival Based ATR", IS&T/SPIE Symposium on Electrical Imaging, January 1996, San Jose, CA.

5. Cyganski, D., B. King, R.F. Vaz, and J.A. Orr, "ROC Analysis of ATR from SAR images using a Model-Based Recognizer Incorporating Pose Information", SPIE 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics, April 1995, Orlando, Florida.

6. R.F. Vaz, D. Cyganski, and B.M. King, "An ROC Comparison of Pose Invariant and Pose Dependent Model Based ATR", 4th ATR Systems and Technology Symposium, Monterey, California, November, 1994.

7. J. Byrne, D. Cyganski, R.F. Vaz, "Implementation and evaluation of an N-Dimensional Coherent Wave DOA estimator", Seventh Signal Processing Workshop, Quebec City, Quebec, Canada, June 1994.

8. C.R. Wright, R.F. Vaz, and D. Cyganski, "Establishing Identity and Pose of Objects From a Library Using Reciprocal Basis Sets and Direction of Arrival Techniques," Proceedings of the SPIE Symposium on Intelligent Robotic Systems, November 1993, Boston.

7 Personnel Supported and Degrees Awarded

1. Faculty
   (a) Prof. David Cyganski, ECE Department, WPI
   (b) Prof. Richard Vaz, ECE Department, WPI
   (c) Prof. John A. Orr, ECE Department, WPI

2. Graduate Students
   (a) Charles R. Wright, Received Ph.D. E.E. Degree in May 1994, Thesis titled: "Multidimensional Direction of Arrival Performance Bounds and Optimization for Nonstationary Noise."
   (b) David Byrne, Received M.S. E.E. Degree in May 1994, Thesis titled: "Implementation and Evaluation of N-Dimensional Coherent Wave DOA Estimators."
   (c) David Chan, Received M.S. E.E. Degree in May 1995, Thesis titled: "Implementation and Evaluation of a Parallel Singular Value Decomposition Algorithm for Direction of Arrival Analysis."

(d) Brian King, Received M.S. E.E. Degree in May 1995, Thesis titled: "Optimization of the LSD/DOA ATR Technique for Non-Stationary Correlated Gaussian Noise."

(e) Barry Hill, Received M.S. E.E. Degree in May 1996. Thesis titled: "Extension of the LSD/DOA ATR System for Multi-Target Discrimination and Recognition of Obscured Targets."

(f) Daniel Cohen, Received M.S. E.E. Degree in May 1996. Thesis titled: "Optimization of a 2-DOF LSD/DOA ATR System."

(g) Witold Jachimcyzk, M.S. E.E. Degree in May 1997. Thesis titled: "Enhancements of Pose-tagged Partial Evidence Fusion for Automatic Target Recognition."

(h) Debra Fraser, M.S. E.E. Candidate.

(i) Jim Kilian, Ph.D. E.E. Candidate.

8 Inventions

None.

References

[1] D. Cyganski, R.F. Vaz, and C.R. Wright, "Model-Based 3-D Object Pose Estimation from linear image decomposition and direction of arrival techniques," *Proc. SPIE Proceedings, Conference on Model-Based Vision, vol. 1827*, November 1992, Boston.

[2] David Cyganski, James Kilian, Deborah Fraser, "Baseline Performance Analysis of the LSD/DOA ATR Against MSTAR Data", SPIE AeroSense '98, Conference 3370, Algorithms for Synthetic Aperture Radar Imagery V, April 1998, Orlando, Florida.

[3] Witold Jachimczyk, David Cyganski , "Enhancements of Pose-tagged Partial Evidence Fusion SAR ATR," AeroSense '97, Conference 3070, Algorithms for Synthetic Aperture Radar Imagery IV, April 1997, Orlando, Florida.

[4] B.K. Hill, D. Cyganski, R.F. Vaz , "SAR ATR via pose-tagged partial evidence fusion," AeroSense '96, Conf. 2757, Algorithms for Synthetic Aperture Radar Imagery III, April 1996, Orlando, Fla.

[5] Hill, B.K., Cyganski, D., R.F. Vaz, "Multi-Target Discrimination with Linear Signal Decomposition/Direction of Arrival Based ATR", IS&T/SPIE Symposium on Electrical Imaging, January 1996, San Jose, CA.

[6] Cyganski, D., B. King, R.F. Vaz, and J.A. Orr, "ROC Analysis of ATR from SAR images using a Model-Based Recognizer Incorporating Pose Information", SPIE 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics, April 1995, Orlando, Florida.

[7] R.F. Vaz, D. Cyganski, and B.M. King, "An ROC Comparison of Pose Invariant and Pose Dependent Model Based ATR", 4th ATR Systems and Technology Symposium, Monterey, California, November, 1994.

[8] J. Byrne, D. Cyganski, R.F. Vaz, "Implementation and evaluation of an N-Dimensional Coherent Wave DOA estimator", Seventh Signal Processing Workshop, Quebec City, Quebec, Canada, June 1994.

[9] C.R. Wright, R.F. Vaz, and D. Cyganski, "Establishing Identity and Pose of Objects From a Library Using Reciprocal Basis Sets and Direction of Arrival Techniques," Proceedings of the SPIE Symposium on Intelligent Robotic Systems, November 1993, Boston.

[10] Charles R. Wright, WPI Ph.D. E.E. Thesis titled:"Multidimensional Direction of Arrival Performance Bounds and Optimization for Nonstationary Noise," May 1994.

[11] David Byrne, WPI M.S. E.E. Thesis titled:"Implementation and Evaluation of N-Dimensional Coherent Wave DOA Estimators," May 1994.

[12] David Chan, WPI M.S. E.E. Thesis titled:"Implementation and Evaluation of a Parallel Singular Value Decomposition Algorithm for Direction of Arrival Analysis," May 1995.

[13] Brian King, M.S. E.E. Thesis titled: "Optimization of the LSD/DOA ATR Technique for Non-Stationary Correlated Gaussian Noise," May 1995.

[14] Barry Hill, M.S. E.E. Thesis titled:"Extension of the LSD/DOA ATR System for Multi-Target Discrimination and Recognition of Obscured Targets," May 1996.

[15] Daniel Cohen, M.S. E.E. Thesis titled:"Optimization of a 2-DOF LSD/DOA ATR System," May 1996.

[16] Witold Jachimcyzk, M.S. E.E. Thesis titled: "Enhancements of Pose-tagged Partial Evidence Fusion for Automatic Target Recognition," May 1997.

[17] Novak, L.M., M.B. Sechtin, "On the Performance of Linear and Quadratic Classifiers in Radar Target Detection," IEEE 1987

[18] Novak, L.M., G.J. Owirka, and C.M. Netishen, "Performance of a High-Resolution Polarimetric SAR Automatic Target Recognition System," The Lincoln Laboratory Journal, volume 6, number 1, 1993

[19] S. Kay, "A Fast and Accurate Single Frequency Estimator", IEEE Trans. on Acoustics, Speech, & Signal Processing, vol. 37, 1989, pp. 1987-1990.

[20] G. Lank and I. Reed and G. Pollon, "A semicoherent detection and Doppler estimation statistic", IEEE Trans. on Aerospace Electronics Systems, vol. 9, 1973, pp. 151-165.

[21] D. Cyganski and D. Fraser, "A Two-Stage Regression-Based Frequency Estimator", *Submitted to:* IEEE Signal Processing Letters.

[22] E. Keydel and S. Lee and J. Moore, "MSTAR Extended Operating Conditions: A Tutorial", Proc. SPIE: Algorithms for Synthetic Aperture Radar Imagery III, 1996, vol. 2757, pp. 228-242.

[23] T. Ross and L. Westerkamp and E. Zelnio and T. Burns, "Extensibility and other model-based ATR evaluation concepts", Proc. SPIE: Algorithms for Synthetic Aperture Radar Imagery IV, 1997, vol. 3070, pp. 213-222.